DOCUMENT RESUME

ED 099 125 PS 007 601

AUTHOR McDaniel, Ernest
TITLE Design of a Longitudinal Study of Educational Effects.
SPONS AGENCY National Center for Educational Statistics (DHEW/OE), Washington, D.C.
PUB DATE Sep 74
CONTRACT OEC-0-72-5283
NOTE 21p.; Paper presented at the Annual Meeting of the American Psychological Association (82nd, New Orleans, Louisiana, September 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS Cognitive Development; *Data Analysis; Data Collection; *Educational Research; *Elementary School Students; Emotional Development; Family Influence; Family School Relationship; Growth Patterns; *Longitudinal Studies; *Research Design; Sampling; Social Development; Socioeconomic Influences; Standardized Tests; Teacher Influence

ABSTRACT

This paper presents the research design for a longitudinal study of elementary school students, which will examine patterns of cognitive, affective, and social growth and identify the school and home variables which influence such growth. Shifts in dependent variables during a single school year will be examined during the initial year for all grades at the elementary school level. During this year, considerable attention will be paid to the actual classroom behavior of the teacher. During the subsequent 2 years, continuing data will be collected for a first and fourth grade cohort. Thus, in a 3-year period a picture of growth during the initial 6 years of schooling will be obtained. Various aspects of the longitudinal study which are discussed in some detail in this research design are: assessment of variables, testing instruments, sampling plan, testing sequence, testing schedule, migrating children, and plans for data analysis. (Author/CS)

ED 099125

BEST COPY AVAILABLE

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Design of a Longitudinal Study of Educational Effects

Ernest McDaniel

Purdue University

A longitudinal study is something like a space shot: It is exceedingly expensive, requires exquisite planning and organization, carries a sharply limited payload, and is entered into with fear and trembling on the part of the participants. Such considerations have induced an appropriate hesitancy in embarking on longitudinal studies. At the same time, it is evident that the most pressing problems of education are more likely to yield to long range, programatic research than to short term status studies.

The design of the longitudinal study presented here has its roots in a 1968 survey of educational policy makers and researchers conducted by the National Center for Educational Statistics to determine research needs. This survey was followed by many conferences and imput from all segments of the educational and research community. After many meetings, several points became relatively clear:

1. Professional people want to know what makes differences in education, i.e., they want to know what will happen to Tommy, who may be black and disadvantaged, when he is taught with the SRA reading program by a teacher with attributes like Miss Wilson.

2. Professional people want to know more about the psychological impact of schooling, i.e., while Tommy is learning to read what is happening to his attitude toward school, his self concept and his personal and social development.

3. Professional people want their specialities represented in any data bank assembled so that they and their colleagues can tease out answers to questions ranging from the effectiveness of special education to the influence of educational technology on learning outcomes.

In 1970, the difficult task of building a preliminary design representing the diverse interests of the educational community, and of developing prototype instruments was assigned to the American Institute of Research in Palo Alto. The magnitude of their tasks is indicated by the report that the preliminary work produced a list of over a thousand variables. Further work reduced the number of variables to 378. A design was proposed which specified testing in grades 1, 3, 6 and 8; and more than twenty-five prototype tests were written.

A paper presented as part of a symposium: "Large Scale Studies of Educational Effects" at the 82nd Annual Meeting of the American Psychological Association, New Orleans, September 2, 1974.

The work reported above was performed under Contract OEC 0-725283, National Center for Educational Statistics, United States Office of Education.

PS 007801

BEST COPY AVAILABLE

In 1972 a contract was awarded to the Purdue Educational Research Center for further elaboration of the design and development and selection of tests for use in the study. Intensive study of the evolving design by the Purdue group, personnel of the National Center of Educational Statistics, and a number of outside consultants suggested several significant shifts in orientation:

1. The luxury of satisfying many segments of the educational community would have to be abandoned in favor of a more sharply focused and readily attainable study.

2. Direct observation of teaching behavior would be an essential aspect of any study purporting to examine educational effects.

3. The eight year longitudinal study might better be conducted as a series of shorter overlapping studies.

4. A sampling plan would have to emerge which would accommodate a compact field force and permit follow-up of students without prohibitive cost.

The design of a longitudinal study of elementary school effects presented in the following pages reflects these orientations. The focus of the study is fairly limited, concentrating on selected home and classroom variables which impinge on cognitive and affective development. Shifts in dependent variables during a single school year will be examined during the initial year for all grades at the elementary school level. During this year, considerable attention will be paid to the actual classroom behavior of the teacher. During the subsequent two years, continuing data will be collected for a first and fourth grade cohort. Thus, in a three year period a picture of growth during the initial six years of schooling will be obtained. Finally, the sampling plan makes use of a high degree of geographic clustering, thus simplifying data collection and follow-up of students.

BEST COPY AVAILABLE

Design of the Longitudinal Study of Elementary School Effects

This study is designed to describe patterns of cognitive, affective and social growth among elementary school children and to identify the school and home variables which influence such growth. The study will seek answers to two kinds of questions: How do children grow and develop during the elementary school years? What are the school and home influences affecting this growth?

Children in grades 1 through 6 will be tested in the fall and spring of the initial year of the study. Intensive data will be collected about the home background of the children and the characteristics and behavior of the classroom teacher. During the two subsequent years, data will be collected from a limited sample of children who originally were tested in grades 1 and 4. By linking the three year segment of growth from the first grade group and the three year segment from the fourth grade group, a picture of development spanning the six elementary school years will be obtained. Figure 1 provides a graphic representation of the design of the study.

Data collected during the initial year will provide for description and analysis of the home and school variables related to achievement and personal-social development. In addition, cross-sectional comparisons can be made between achievement and personal-social development of children at six different grade levels. For these studies, the classroom group will be the unit of analysis.

During subsequent years the study will concentrate on longitudinal analysis. It will be concerned with stability and change in patterns of school achievement, affective development and social growth. For these studies, the individual child will be the unit of analysis.

The study should help answer a wide variety of questions related to child development and instructional processes. It will permit the tracing of developmental patterns for children in general and for special subgroups, such as black, poor, handicapped, gifted, inner city or rural children. It will permit investigation of the impact of various teaching styles and instructional strategies. It will lead to a more complete understanding of the role played by parental attitudes in the child's achievement, attitude toward school, self-concept and social development. Finally, the study will offer opportunities to investigate the complex interactions among home, school and student variables as they evolve through the elementary school years.

Variables

The variables chosen for investigation in this study were selected from a wide range of potential variables describing the child, the home and the school. The variables that describe the child are the dependent variables. Those describing the home and school are the independent variables. Some characteristics of the child, the home and the school function as control or classification variables.

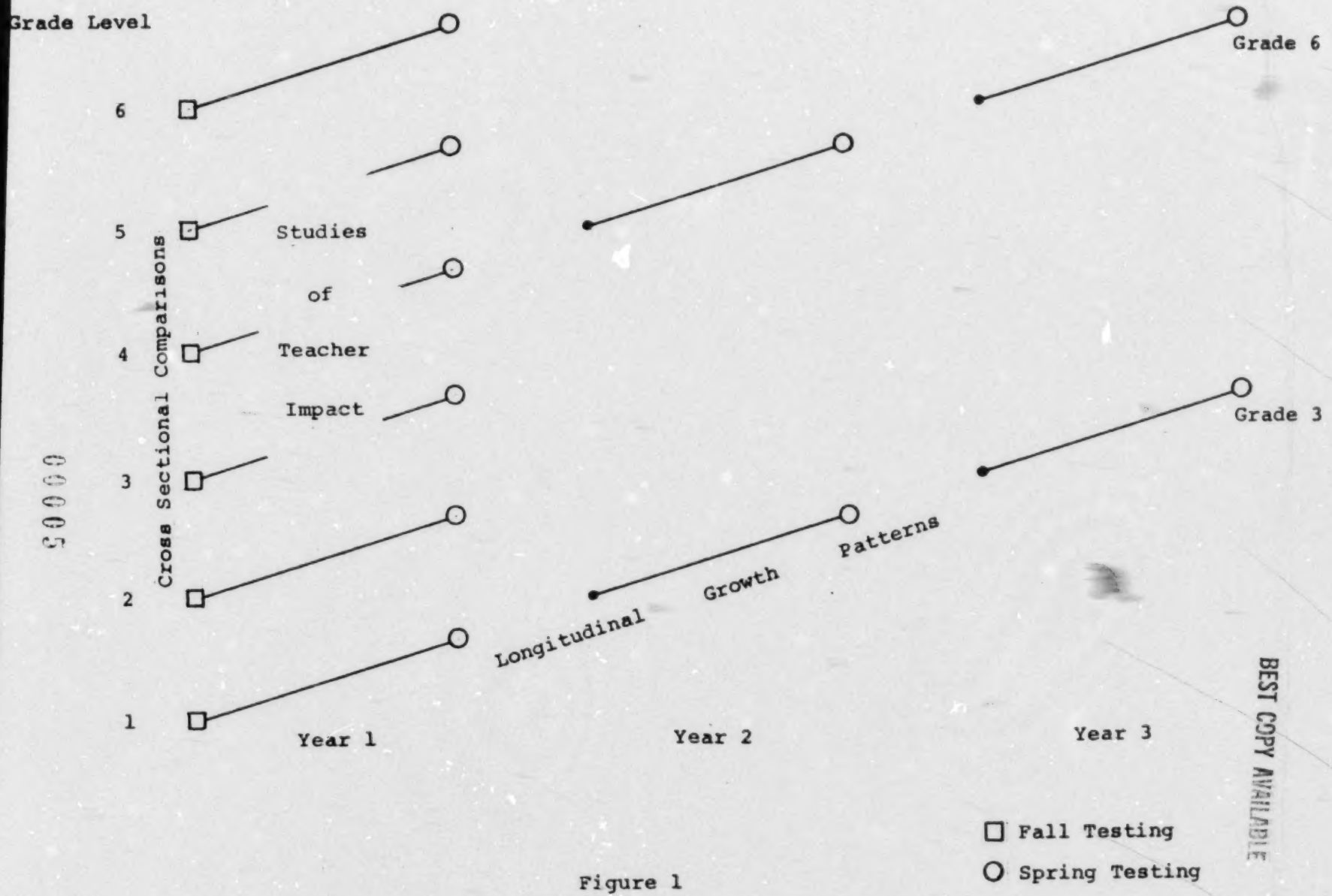

Figure 1

Design of Study

50000

BEST COPY AVAILABLE

BEST COPY AVAILABLE

Dependent variables. The following list of dependent variables represents a stringent sampling in the areas of achievement, cognitive processes and affective development.

Reading achievement
Arithmetic achievement
Problem solving
Concept formation
Attitude toward school
Self-concept
Personal-social development

In the achievement area, reading is the key variable. Research literature documents its pervasive relationship with academic achievement in all curriculum areas. Arithmetic achievement represents the quantitative dimension.

In the cognitive area, problem solving and concept formation are the dependent variables. These two variables reflect the adaptive, coping strategies of the developing child.

In the area of affective development, attitude toward school, self-concept and personal-social development were selected as representative variables.

Independent variables. The major independent variables in a study of elementary school effects must reflect the distinctive features of school and home which contribute to the growth of the child. The following list of independent variables represents a selection from the school variables that the literature suggests may be important.

School organizational climate
Classroom physical environment
Classroom materials and equipment
Size of class
Reading strategies
Arithmetic strategies
Verbal facility of teacher
Teaching behavior
- Warmth
- Enthusiasm
- Clarity
- Variety
- Individualization
- Feedback
- Cognitive demand
- Freedom
- On-task Activity

The school organizational climate sets the context within which teaching takes place. Teaching is further constrained by the physical environment of the classroom, the adequacy of materials and equipment and the number of pupils in the class. Given the context, the teacher is certainly the

BEST COPY AVAILABLE

most important element in the classroom. The independent variables stress the teacher's instructional strategies, his verbal facility and a selected set of specific teaching behaviors.

The literature also suggests a multitude of potential variables related to the home. The following list of independent variables has been chosen to reflect important parental attitudes.

Parental aspiration for child
Learning environment
Concern for education
Maturity demand
Quality of TV
Parental restriction of TV
Activities of parents
Parental attitude toward education
Acceptance of child
Permissiveness
Restrictiveness
Parental attitude toward child nonconformity
Parent's authoritarian attitudes

Both the child's achievement and attitude toward school are likely to be influenced by such variables as the richness of the learning environment in the home, parental aspirations for the child, and concern for education. The child's self-concept and personal-social development may reflect maturity demand in the home, acceptance of the child, and the permissiveness or restrictiveness of the parents.

Control or classification variables. Information is collected which will permit the control of some variables and the classification of children, homes, teachers or schools on the basis of important characteristics. The major control or classification variables are listed below.

Demographic information about child
General ability
Pre-school education
Demographic information about home and family
Demographic information about teacher
Teacher satisfaction

This presentation of the important variables in the study omits many sub-scores and individual items. Also omitted are descriptions of supplementary data collected on the limited sample of children who will be followed for three years.

Relationships among variables. The variables are introduced as clusters of variables describing the child, the home and the school in order to make explicit the structure of the study. The study is designed to permit analysis of the network of relationships among the variables.

BEST COPY AVAILABLE

If the child's attitude toward school is an important dependent variable, then both the parents' attitude toward education and the teacher's warmth and enthusiasm become relevant independent variables. Figure 2 illustrates the interactive nature of the home and school variables impinging on the child.

Multimethod Assessment of Variables

In two important areas, the personal-social development of the child and the descriptions of teaching behavior, multimethod approaches to measuring the variables have been built into the instrumentation of the study. The primary measure of personal-social development is a sociometric instrument on which pupils rate their peers on a variety of behaviors. In addition, first and fourth grade teachers provide similar information about the small group of children selected to be followed for three years. If positive correlations are obtained from these interlocking measures, such correlations will offer support for the validity of the instruments.

Multimethod approaches also have been employed in the description of teaching behavior. Pupils describe the teacher's instructional behavior. Teachers complete a form describing their own behavior and trained observers rate teachers on the same dimensions. Again, positive correlations among these three measures would offer support for the validity of the instruments used to assess teaching behavior.

Instruments

Fifteen instruments were selected or developed to measure the variables in the study. These instruments are listed and described below.

1. Class Roster - a checklist form that obtains from the teacher a complete class list and preliminary information about each child.

2. Raven's Coloured Progressive Matrices - a non-verbal test measuring general ability of children in grades 1 through 6.

3. Clymer-Barrett Prereading Battery - a readiness test based on recognition of letters and sounds. This test will be used at the beginning of grade 1.

4. Stanford Early School Achievement Test (Mathematics) - a test to measure number concepts. This test will be used at the beginning of grade 1.

5. Stanford Achievement Tests, 1973 edition (Reading and Mathematics) - tests designed to assess reading comprehension and mathematics computation and concepts.

6. Purdue Elementary Problem-Solving Inventory (Picture-Book Form) - a problem-solving test employing cartoons depicting children in problem situations for which solutions are selected.

BEST COPY AVAILABLE

7. Purdue Concept Formation Test - Conservation - a test designed to measure the ability to conserve quantity, length, mass, area and volume.

8. Attitude Toward School - a questionnaire designed to assess attitude toward school in general, attitude toward schoolwork and attitude toward teacher.

9. Piers-Harris Children's Self-Concept Scale - a self report indicating how a child feels about himself used for grades 4 to 6. A shortened form adapted from Piers-Harris is used for grades 1 to 3.

10. Peer Ratings of Personal-Social Development - sociometric scales obtaining ratings of each child by three other children on 12 characteristics.

11. Pupil Information Booklet - a booklet completed by the teacher for each child in the cohorts. The booklet provides demographic information and ratings of the child's personal-social development.

12. Parent Questionnaire and Parent Interview Schedule - a questionnaire obtaining demographic information from the parents and information related to educational aspirations and parent-child relations. The interview schedule is a modification of the questionnaire.

13. Pupil Description of Teaching - a pupil report which describes the teacher's instructional practices.

14. Teacher Questionnaire - a self report which obtains demographic information, and measures of organizational climate, instructional practices and verbal facility.

15. Teacher Observation Scales - observational procedures for recording classroom descriptions, instructional strategies and teaching behaviors. Teaching behaviors are observed on dimensions which parallel those on the Pupil Description of Teaching and the Teacher Questionnaire.

Sampling Plan

Several principles have guided the sampling plan. The population defined for sampling should be as large as possible. The sampling plan should enable any school within population regions and strata to have an equal chance of being drawn. For economy and close field coordination and supervision of the data collection, the sample should have high geographic clustering. To expedite and standardize procedures and conditions at each site, the sampling should be simple and make use of sampling frames that can be constructed locally on the basis of data that can be expected to be available at each site. Lastly, the sampling plan should be flexible enough to scale up or down according to the total funds allocated to the study and also to meet field contingencies such as uncooperative school

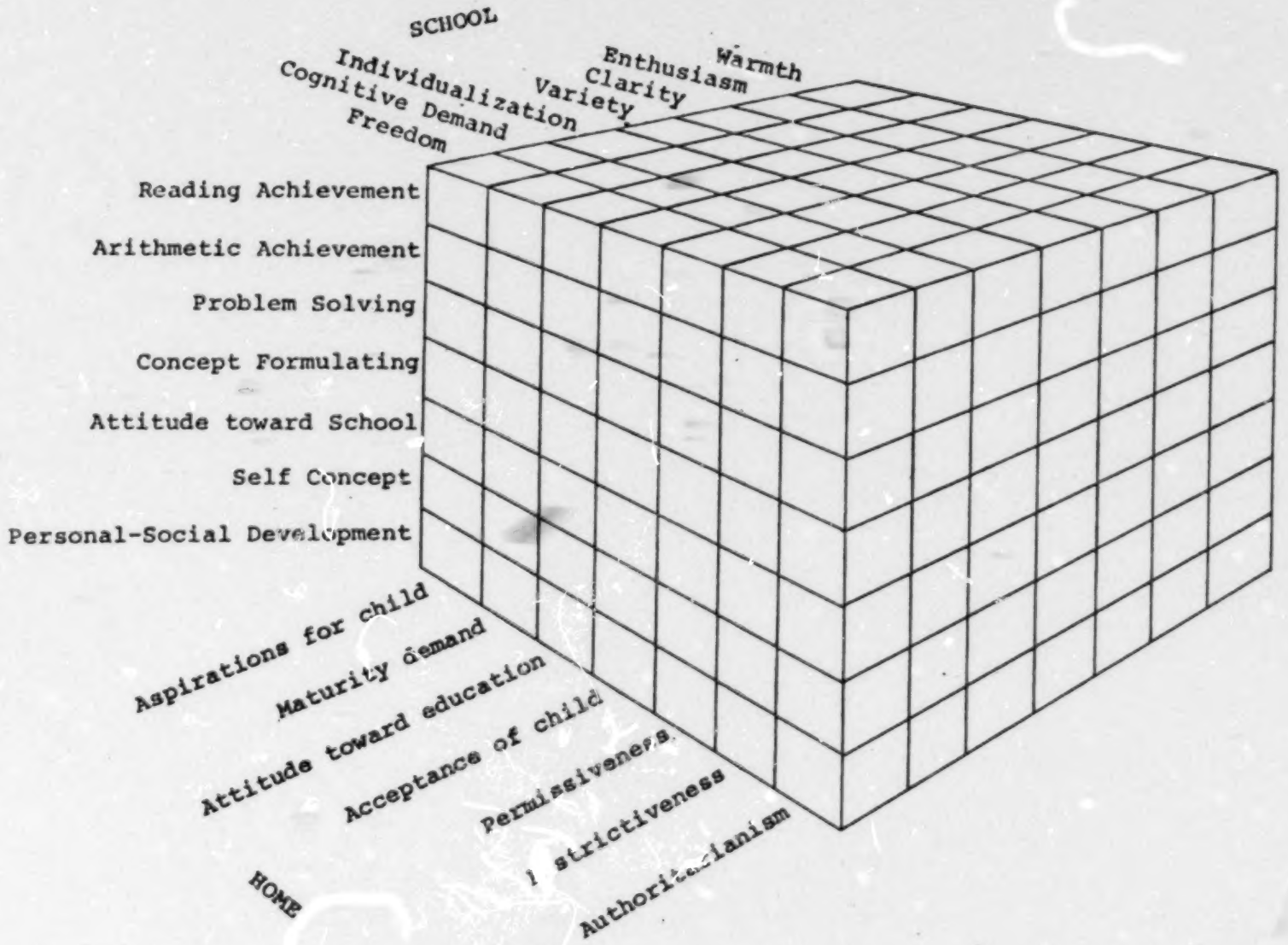

Figure 1

Interactive Nature of Variables

BEST COPY AVAILABLE

BEST COPY AVAILABLE

districts.

The sampling frame will consist of the 243 Standard Metropolitan Statistical Areas which contained in 1970 about 139 million persons or about 70 percent of the total population in the United States, of which 120 million were white and 17 million were black. Since the SMSA's only contain about 10 percent of the U. S. land area, this population affords a high degree of geographic clustering. The Bureau of the Census defines a SMSA as an integrated social and economic unit with a large population nucleus. Each SMSA contains a central city of 50,000 or more inhabitants, or two cities with contiguous boundaries forming a socio-economic unit of 50,000, the smaller of which must have a population of at least 15,000. The 243 SMSA's will be sorted into four geographic regions, the Northeast, Midwest, Far West, and South, and two SMSA's will be randomly selected from each region, yielding 8 basic sites.

A Site Manager and two Survey Team Leaders will be recruited from local universities or federally-sponsored research agencies at each site. The main responsibilities of the Site Manager will be to gather data to carry out the sampling within the site, obtain the cooperation of the local educational authorities, and direct the data collection in the schools. To gather the data within a period of four weeks during the fall and spring testing periods, each of the two Survey Team Leaders will assume responsibility for leading teams in surveying six schools. The schools will be surveyed in the same sequence to insure that the time between testings is roughly the same for all schools. A strong effort will be made to recruit minority group members to serve on the staff at each site.

It can be estimated that about three classes will exist at each grade level in the schools selected for the study. Estimating 30 pupils in each class, the sampling procedure will yield about 540 students per school, 6,480 at each site and 51,840 in the total national sample. The following tabulation summarizes the estimates based on testing all classes in the survey schools:

Sampling Level		Cumulative Number of Pupils
1	Class	30
3	Classes at each grade level	90
6	Grade levels at each school	540
12	Schools at each site	6,480
8	Sites	51,840

Rounding down, the above figures define the initial year sample as approximately 51,000 students and 1,700 teachers.

BEST COPY AVAILABLE

Eight children will be sampled from all first and fourth grade classes to form the cohorts for further longitudinal analysis. The following tabulation presents estimates of the sample size for the second and third years of the study.

Sampling Level		Cumulative Number of Pupils
1	Subsample from each class	8
3	Classes at each grade level	24
2	Grade levels at each school	48
12	Schools at each site	576
8	Sites	4,608

These figures indicate that 4,600 children will be in the study after the initial year, approximately 2,300 children in each cohort.

Educators and federal program managers are concerned with poor, minority children in the central cities, and these major target groups should be oversampled. Thus, for each central city, a list of public schools will be obtained with an indication of each school's eligibility for Title I funds under ESEA. The central staff of the Chicago Public Schools has surveyed the 15 largest public school systems in the United States and found that all have some objective basis for determining eligibility such as census data, percent of children in families receiving Aid to Families of Dependent Children support, or local surveys. The schools within the district will be rank-ordered according to the objective criterion available, and eight schools will be randomly selected from the half of the list containing schools in disadvantaged areas; two schools will be chosen randomly from the rest of the list. In addition, one suburb, village, or town within the SMSA outside the central city will be randomly chosen, and two schools within the public system will be chosen randomly for inclusion in the survey.

The proposed sampling plan can be modified. If economies become necessary, limiting the study to one site per region with a total of four sites would cut the costs considerably, though not in half since the costs of printing tests and processing the data are lower on a per unit basis with larger samples. Also, practical considerations may require modification of the sampling plan. Coleman, when collecting data for the Equality of Educational Opportunity survey, was unable to obtain access to large city school systems in several cases. Therefore, it may become necessary to select the next largest municipality in one or two SMSA's.

Testing Sequence

The data collecting process will proceed through a series of sequential steps:

BEST COPY AVAILABLE

1. Each teacher will complete a class roster form. This form provides the master record of children in the study. The form also will permit the selection by the project staff of eight children from each class for the cohort sample.

2. A battery of achievement, cognitive, self-concept, attitude toward school and socialization tests will be administered during a four week fall testing period.

3. During the winter months, teachers will complete a questionnaire about their own teaching practices and be observed three times by trained observers. Children will take home a questionnaire to be filled out by a parent. Parents who do not return the questionnaire will be interviewed. Teachers also will complete an information booklet for each child selected for the cohorts.

4. During a four week period in the spring, the battery of tests will be readministered, with the exception of the test of general ability. Children also will complete a form describing the teacher's instructional practices.

5. During subsequent years, the fall and spring battery of tests will be administered to all children in the cohorts. In the spring, teachers also will complete an information booklet for each child in the cohorts.

6. No new parental information will be collected during subsequent years and no further assessments of teaching behavior will be obtained.

7. During the initial year, testing will include all children enrolled in a project school during the fall and spring testing periods.

8. All children in the cohort groups migrating out of the original schools, but remaining within the boundaries of a site, will be retained in the study over the three year period.

9. A 20 percent sample of cohorts migrating out of the original site will be obtained and the data analyzed to provide information about biasing effects of selective migration of students.

Testing Schedule

The number of tests required of the children in this study is quite large. In addition, a number of variables reflect attitudes and feelings that are likely to start changing in response to the new school environment from the very first days in class. It seems desirable, therefore, to initiate testing as soon as the children have made an initial adjustment to school. The schedule provides for three major data gathering periods, fall, winter and spring. The fall and spring testing periods provide measures of the major dependent variables. The winter period provides data on the home and school. The fall testing period should start the last

BEST COPY AVAILABLE

week of September, the winter period in mid-January, and the spring period in mid-March.

The fall and spring testing periods are four weeks long. The schedule for first grade testing, presented below, illustrates the collection of attitude and self-concept data in two separate sessions during the first week, and the collection of achievement measures in three sessions the next week. Problem-solving and concept formation tests will be administered in two sessions in the third week. During the fourth week, the Raven's Coloured Progressive Matrices will be administered in one session and peer ratings will be collected in another session. No single testing session exceeds 45 minutes. Trials with the instruments indicate that the time estimates are generous. Testing at other grade levels will follow a similar schedule.

First Grade Schedule

Fall

First Week	
Attitude Toward School	20 minutes
Self-Concept	40 minutes
Second Week	
Clymer-Barrett Prereading Battery	
Part A	18 minutes
Part B	15 minutes
Stanford Early School Achievement Test (Math)	20 minutes
Third Week	
Purdue Elementary Problem-solving Inventory	40 minutes
Purdue Concept Formation Test - Conservation	40 minutes
Fourth Week	
Raven's Coloured Progressive Matrices	30 minutes
Peer Ratings	40 minutes

Spring

First Week	
Attitude Toward School	20 minutes
Self-Concept	40 minutes
Second Week	
Stanford Achievement Tests	
Reading	
Vocabulary	20 minutes
Reading Comprehension, Part A	20 minutes
Reading Comprehension, Part B	20 minutes

BEST COPY AVAILABLE

Spring (continued)

Second Week (continued)

Stanford Achievement Tests (continued)

Arithmetic	
Concepts	25 minutes
Computation	30 minutes

Third Week

Purdue Elementary Problem-solving Inventory	40 minutes
Purdue Concept Formation Test - Conservation	40 minutes

Fourth Week

Pupil Description of Teaching	30 minutes
Peer Ratings	40 minutes

Teachers and Parents. Data will be collected from teachers and parents primarily during the winter. The time estimates for instruments are listed below.

Class Roster	60 minutes
Pupil Information Booklet (30 minutes per child, eight children)	240 minutes
Teacher Questionnaire	60 minutes
Parent Questionnaire	60 minutes
Parent Interview	60-90 minutes

Migrating Children

One of the major concerns in planning the longitudinal study has been the problem of migrating children. Although nationally perhaps 5 to 10 percent of the children move out of a school's attendance district during a school year, mobility rates in inner-city schools runs as high as 50 percent. Migrating children do not, however, constitute a problem during the initial year of the study. During this year, primary data analysis will center around classroom means. Even in areas of high mobility, classroom replacements share many demographic characteristics with children who leave. Suburban children migrate through suburban schools and inner-city children migrate through inner-city schools. For this reason, mean scores of class groups obtained in the fall and spring should not be greatly affected by migration alone.

The migration of cohorts studied over a three year span, however, could be troublesome. Concern has been expressed that cohorts in a relatively compact set of classes during the initial year will be found dispersed over a great many classes during the following years.

BEST COPY AVAILABLE

The problem of proliferating classrooms may be considerably smaller than anticipated. Typically the number of first grade classes in an elementary school defines the maximum number of classes in each of the upper grades. Thus, if a school with three first grade classes participates in the study then, discounting migration, all children making normal progress will be found in three second grade classes of the same school during the second year of the study.

The problem of cohorts who migrate from the original school is somewhat larger. Two procedures, however, should reduce the problems of locating children who change schools. First, the school record folders of all children selected as cohorts for the study will be flagged with a gummed label displaying a colored insignia identifying the folder as belonging to a child participating in the National Longitudinal Study of Elementary School Effects. In the folder will be a self addressed postcard, identified with the same colored insignia, requesting school personnel to mail the card giving new school address in the event of transfer of records.

The second procedure designed to facilitate tracing students who move is the collection of special information from parents. On the parent questionnaire and interview schedule, there is a request for the name and address of a close friend or relative who will have a forwarding address in case the respondent moves. This information together with that supplied by the school personnel should enable accurate tracing of a high percentage of the students who change schools.

Once the children are located, a second problem is that of imposing a lengthy testing schedule on a single child enrolled in a school which is not participating in the study. This latter problem can be approached through test sampling procedures. Migrating children will be administered a subset of three tests. For example, some first grade children in the cohort group who changed schools between fall and spring testing might be administered Reading Achievement, Self-concept, and problem solving. Other migrating first grade children might receive the Math Achievement, Attitude Toward School, and Concept Formation. Peer ratings will be obtained by drawing a subsample of the child's new classmates. All testing of migrating children would be accomplished by special examiners from the project staff. No assessments of teaching behavior will be obtained for the new class.

All cohorts changing schools within a site will be traced and tested by the resident project staff. The cost of tracing and testing individual cohorts migrating out of the site area will be very high. A sample of 20 percent of these children will be traced and tested using test sampling procedures. The data obtained from the small subsample of migrating cohorts will be used to identify and describe the nature of any biases introduced by selective migration.

BEST COPY AVAILABLE

Data Analysis

Plans for analyzing the data must take account of the unique relationships among variables and among patterns of variables mentioned above. Account must also be taken of the kinds of scales used and of the large number of variables to be handled.

Thus, the statistical analysis of the data will proceed in stages, each stage designed to simplify the analysis in the next stage as much as possible, while still preserving a maximum of information about relationships among the variables.

The first stage of data analysis is designed to combine some of the discrete variables into more functional indices; for example, employment of parents, family income, type of dwelling, and number of rooms should combine to form a single index of socioeconomic status.

During stage two of the analysis, simple descriptive statistics (means, standard deviations) and correlations will be generated for major categories of children.

In stage three, factor analysis of dependent and independent variables will be performed in an effort to reduce variables within each set and to combine data collected with several instruments about a single dimension of behavior. The Plowden report (Peaker, 1971) provides an excellent example of data reduction using these procedures. In addition the stability of factor structures across major classification groups will also be examined.

In stage four, multivariate regression techniques will be applied to determine the predictive relationships among the independent and dependent variables.

Each stage of the data analysis plan is further detailed below.

Stage 1. Inventory of Categorical Variables Within the Control Group

Although it would be of great value to be able to perform separate analyses of the relationships among the dependent and independent values for each cell of a crossclassification of individuals by sex, grade, race, SES, etc., a complete cross-classification of this type would involve too many cells (and thus too few observations in each cell). It will be necessary to classify children by grade, and certainly possible to cross-classify by sex and race. However, before attempting to use the other categorical variables (SES, Pre-School Experience, Dwelling Type, etc.) for a more refined classification, preliminary data analysis must be done to see how many children can be observed in each cell of various kinds of cross-classification. Hence, the first step in the analysis of the data will be to construct contingency tables among all of the categorical control variables.

BEST COPY AVAILABLE

Analysis of these tables, perhaps through Lazarsfeld's latent structural analysis (1966), will yield a picture of the relationships among the categorical variables, and will allow a subset of these variables to be chosen for purposes of aggregating the data for separate statistical analyses.

Of the remaining categorical control variables, those that are highly related to the categorical variables used for aggregation of the data will not be used further (since their influence is already reflected by the variables used for aggregation). The other remaining categorical variables will be combined, if possible, into one or more ordinal scales (indices) reflecting socioeconomic status, special talents or handicaps, and home educational stimulus (e.g., pre-school training, home language environment, etc.) of each child. These scales, together with the numerical (ordinal and interval scale) control variables, will be used as conditioning variables (moderator variables) in the remaining steps of the analysis.

Stage II. Data Analysis of Dependent and Independent Variables

All of the dependent and independent variables in the study can be expressed at least on an ordinal scale. Thus, all of these variables can be treated as continuous variables and subjected to standard univariate and multivariate analyses. All of the analyses described below will be done separately for each cell of the cross-classification obtained in Stage I.

(a) Means and variances will be computed for all independent, dependent, and numerical control variables.

(b) Zero-order correlations will be computed among all of the variables named in (a).

Using the output of (b) and either canonical correlation analysis and/or step-wise partial correlation analysis (with the orders in which the variables are entered being determined by a priori consideration of the importance of the variables), attempts will be made to reduce the number of numerical control variables to a more manageable size. Once these numerical control variables are selected, these variables will be used as conditioning variables, and:

(c) Conditional means will be computed for all independent and dependent variables.

(d) Residual means and variances will be computed for all variables.

(e) Partial correlations will be computed among all of the independent and dependent variables.

For the purpose of further analysis, the quantities obtained in (d) and (e) can now be treated as ordinary means, variances, and

zero-order correlations, based on a reduced number of observations (Morrison, 1967; Rao, 1965; Anderson, 1958).

Stage III. Reduction of the Number of Independent and Dependent Variables

As mentioned above, an unique aspect of this study is the use of multiple measurements of common behavioral dimensions. Although this feature of the study should lead to greater accuracy of conclusions and greater flexibility, it does mean that many of the variables, both within the set of independent variables and within the set of dependent variables, are highly correlated (indeed, have a single factor structure). Hence, unless these single factor configurations of variables are reduced and summarized by a single variable, a serious problem of multicollinearity may result, thus complicating interpretations of the relationships among the dependent and independent variables. In consequence, at this stage of the analysis, Procustes factor analysis techniques will be used to obtain a smaller class of dependent and independent variables which will (1) retain all or most of the information about the behavioral dimensions measured by the larger class of variables, and (2) have a direct and namable relationship to the behavioral dimensions measured by the above mentioned clusters of highly dependent (single-factor) variables. Principal components will be obtained separately for the dependent and independent variables. Each of these components will then be rotated in turn through a set of target vectors reflecting the behavioral dimensions deemed common to the variable clusters present in these variable sets. From these factor analyses (one for the dependent and one for the independent variables), all factors common to the variable clusters will be retained, as well as all factors exhibiting significantly large variances. Where several variables load almost entirely on a single factor, only that factor will be retained for the later analysis. However, variables loading lightly on several factors (and/or with a unique and large component of variance) will be retained for the analysis, in place of the factors obtained. In short, the Factor Analysis will be used not to replace all of the variables, but only to identify multicolliinearites in the data and to eliminate these by replacing highly correlated variables by the single factor (or factors) accounting for all of their variation. (If time and available computer facilities permit, however, separate regression analyses relating the dependent factors to the independent factors may be attempted.)

Besides permitting a reduction in the number of variables considered, the factor analysis performed at this stage will be retained for descriptive purposes, since it will indicate the behavioral domains sampled by the study, and give some indication both of the reliability of these domains (in the presence of the control variables), and of the validity of the study (i.e., that the behavioral domains which the study has been designed to sample have indeed been sampled).

BEST COPY AVAILABLE

It should be remarked that these factor analyses will be performed separately for each cell of the cross-classification by control variables obtained in Stage I. At the present time, it is planned that reduction of variables will be done only if the same reduction can be achieved simultaneously in all cells. The requirement, however, may have to be relaxed somewhat if the factor analyses in various cells are not roughly comparable. If it turns out, however, that the factor analyses are not comparable across cells, this result will, of course, be a highly significant finding of the study.

Stage IV. Multivariate Regression

Having reduced the complexity of the analysis as much as possible, in each cell a separate multivariate regression analysis will be performed to determine predictive relationships among the independent and dependent variables, and to test the tentative hypotheses about such relationships stated earlier in this report. It should be remarked that for purposes of this analysis (and for the analysis in previous stages), the fall measurement of any achievement variable which was also measured in the spring will be used as an independent (predictor) variable for the spring measurement. This technique for measuring change (growth) when multiple repeated measurements of the same variable are obtained has been recommended by Bock (1963) and others.

Also, a multivariate analysis of covariance for the dependent variables across cells, with the independent variables as covariates (Gleser & Olkin, 1969, 1972; Rao, 1958, 1959, Pottoff & Roy, 1964) will be used to determine the influence of the categorical control variables on the relationships between the independent and dependent variables. If no significant differences across grade levels are found, the separate regression analyses for these grade levels can be pooled (with grade level as an independent variable) to produce an overall growth curve for the dependent variables over time (grade level). This growth curve can be later compared with growth curves obtained from the longitudinal study.

In addition to overall multivariate growth curve analyses, marginal (univariate) growth curve analyses on each dependent variable will be performed so as to obtain a more detailed picture of what independent variables influence growth in each dependent variable. These more detailed studies will be used to supplement the overall picture provided by the multivariate analysis.

Another, more detailed study, that can be attempted as time and budget permit is a canonical correlation analysis between the dependent and independent variables. This analysis will study predictability of configurations of dependent variables (dependent canonical variates) by prior (antecedant) patterns of independent variables (independent canonical variates). If the canonical

BEST COPY AVAILABLE

variates so obtained coincide roughly with (or have high correlations with) the factors obtained by the analysis of Stage III, this will provide a significant interpretation of the dependencies revealed by the multivariate regression analysis.

Taken together, these analyses can provide a more complete and detailed picture of the effects of home and school environments on the cognitive, affective, and social development of children in the primary schools than have been available from previous studies.

00021